

.conf2013

**YOUR DATA
NO LIMITS**

Time After Time – Comparing Time Ranges in Splunk

Lisa Guinn

Sr Instructor, Splunk

#splunkconf

splunk>

Legal Notices

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC. The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not, be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Splunk Storm, Listen to Your Data, SPL and The Engine for Machine Data are trademarks and registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners.

©2013 Splunk Inc. All rights reserved.

About Me

- Splunk Senior Instructor since 2009
- Frequent contributor to Splunk Answers
- Love Splunk search language puzzles

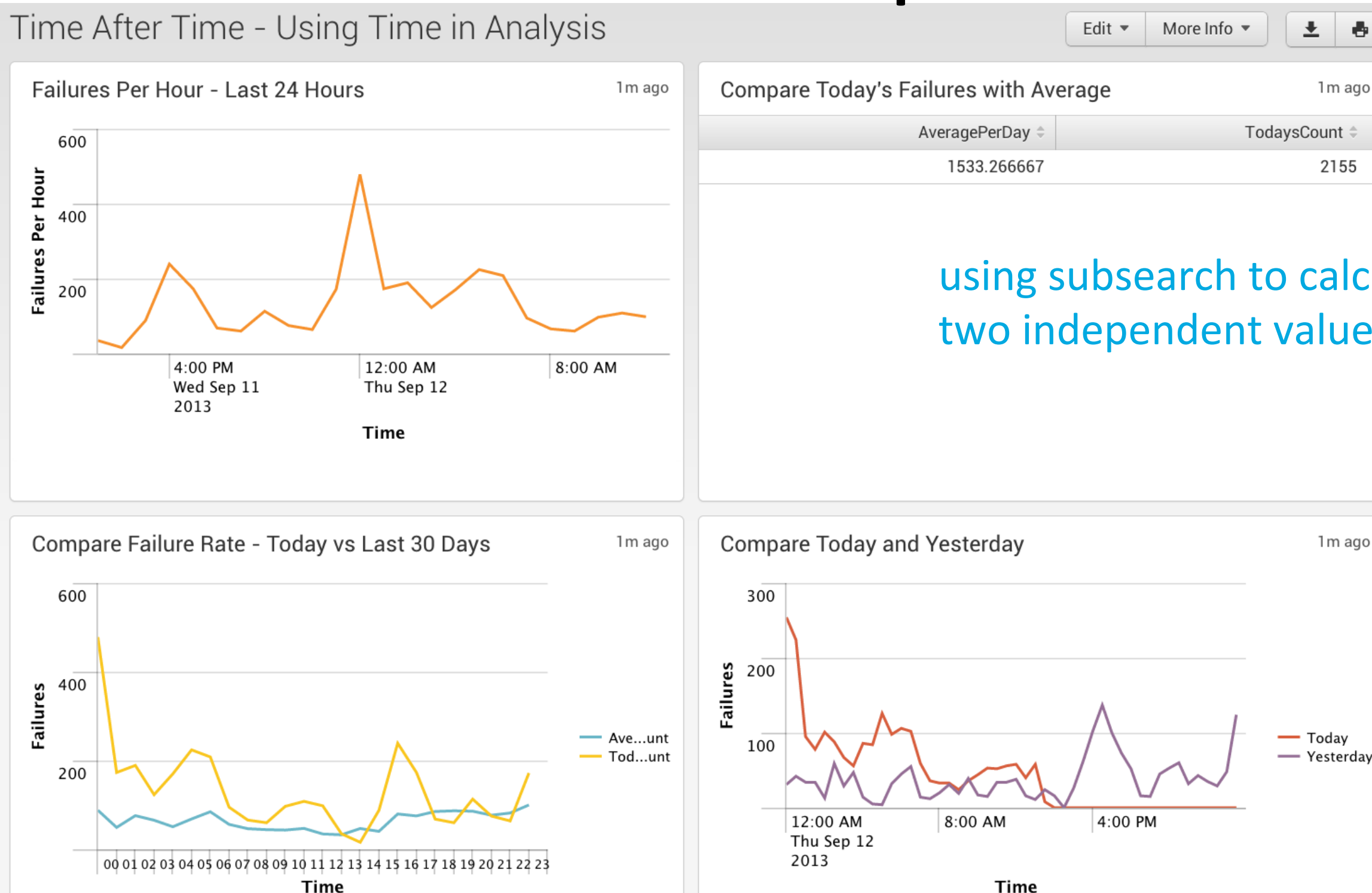
Agenda

Goal: Demonstrate a variety of techniques for creating time-based comparisons

- Review Splunk's Basic Time Fields and Functions
- Examine Some Common Time-based Analysis Techniques
 - Grouping events
 - Computing statistics "across time"
 - Comparing time ranges
- Show Some Tips and Tricks Along the Way
- Answer Your Questions!

Time-based Comparisons

simple
timechart



comparing
different time
ranges

comparing
equal time
ranges

Splunk Time Fields

- Provides context for understanding events
- All events in Splunk are associated with time
- Internal time fields



<code>_time</code>	UTC time based on event timestamp in Unix time format
<code>_indextime</code>	UTC time when event was indexed in Unix time format

`_time` is the field you should use for your time analyses!

Default Datetime Fields

- date_hour
 - date_mday
 - date_minute
 - date_month
 - date_second
 - date_wday
 - date_year
 - date_zone
- These fields come directly from the text of the event;
They do not always exist!
No timezone conversion is applied

87.240.128.18 - - [05/Aug/2013:14:00:53] "POST /product.screen?"

Some Tools to Manipulate Time

- Splunk commands

bucket	Puts time (or numerical values in fields) into discrete sets
timechart	Creates a time-series chart and related table of statistics
eval	Calculates an expression and puts the value into a field

- eval functions

now()	Returns the time that the search started in Unix time
relative_time(t,s)	Returns a new time based on applying the specifier s to time t
strftime(t,f)	Returns a formatted time by applying format f to time t
strptime(ts,f)	Returns Unix time by parsing the time string ts with format f

<http://pubs.opengroup.org/onlinepubs/007904975/functions/strptime.html>

.conf2013

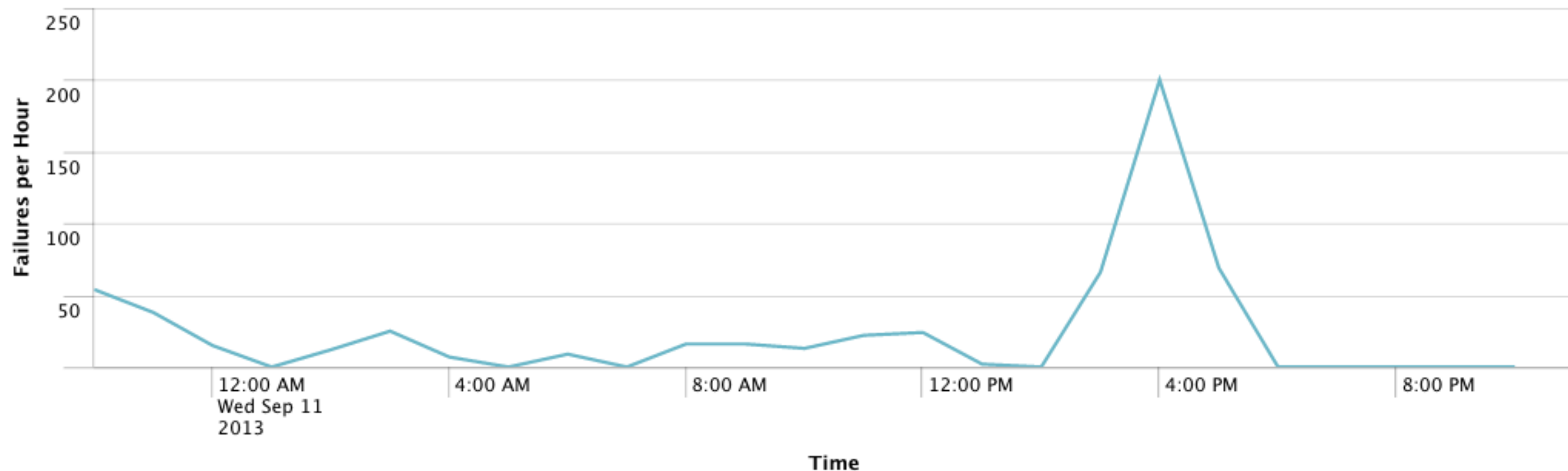
**YOUR DATA
NO LIMITS**

Grouping Events by Time

splunk>

A Common Report: Failures Per Hour for Last 24 Hours

? What is the pattern of failures over the last 24 hours?



Failures Per Hour for Last 24 Hours

? What is the pattern of failures over the last 24 hours?

🔍 `tag=failure earliest=-24h@h latest=@h
| timechart count span=1h`

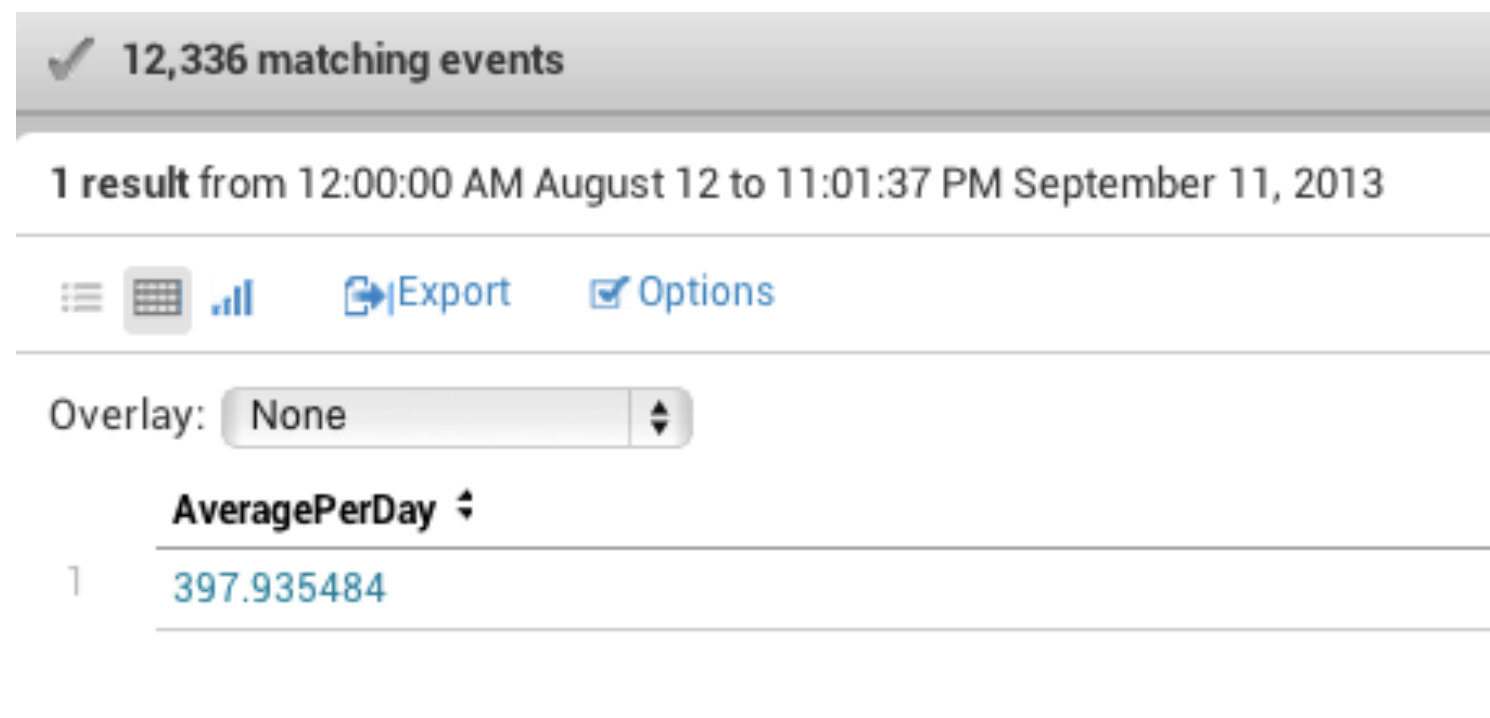
Note: in the slides, I will be showing the time range of each search using the **earliest** and **latest** search terms;
Usually, you could use the green Time Range Picker instead

What Does it Mean?

- What is the Failures per Hour chart showing us?
 - We can see peaks, but is that normal or unusual?
- It would be better if we had a basis for comparison
- Let's start by looking at averages by day...
 - We will come back to the hourly numbers later

Average Failures Per Day Over Last 30 Days

? What is the average number of failures per day?



The screenshot shows a Splunk search interface. At the top, a grey bar indicates '12,336 matching events'. Below this, a text line states '1 result from 12:00:00 AM August 12 to 11:01:37 PM September 11, 2013'. A toolbar contains icons for list, table, and bar chart views, along with 'Export' and 'Options' buttons. An 'Overlay' dropdown menu is set to 'None'. The main results table has a single column header 'AveragePerDay' and one data row with the value '397.935484'.

✓ 12,336 matching events	
1 result from 12:00:00 AM August 12 to 11:01:37 PM September 11, 2013	
⋮ ⌄ 📊 ➦ Export ⌄ Options	
Overlay: None ⌄	
AveragePerDay ⌄	
1	397.935484

After we get the average, we can compare it to today's number...

Average Failures Per Day Over Last 30 Days

? What is the average number of failures per day?

Q tag=failure earliest=-30d@d latest=@d
| bucket _time span=1d ← bucket sets _time to the beginning of the day, "bucketing" the events
| stats count by _time
| stats avg(count) as AveragePerDay ← stats calculates the count for each day

OR

• tag=failure earliest=-30d@d latest=@d
Q timechart span=1d count ← timechart does the same work as bucket + stats above
| stats avg(count) as AveragePerDay

Making the Comparison

	AveragePerDay ↕	TodaysCount ↕
1	402.103448	496

Today's failure count *is* higher than average!

🔍 tag=failure earliest=-30d latest=@d
| timechart span=1d count as dailyCount
| stats avg(dailyCount) as AveragePerDay
| **appendcols** [**search** tag=failure earliest=@d latest=now
| stats count as TodaysCount]

← subsearch counts today's failures, and
appendcols adds the result to the
outer search

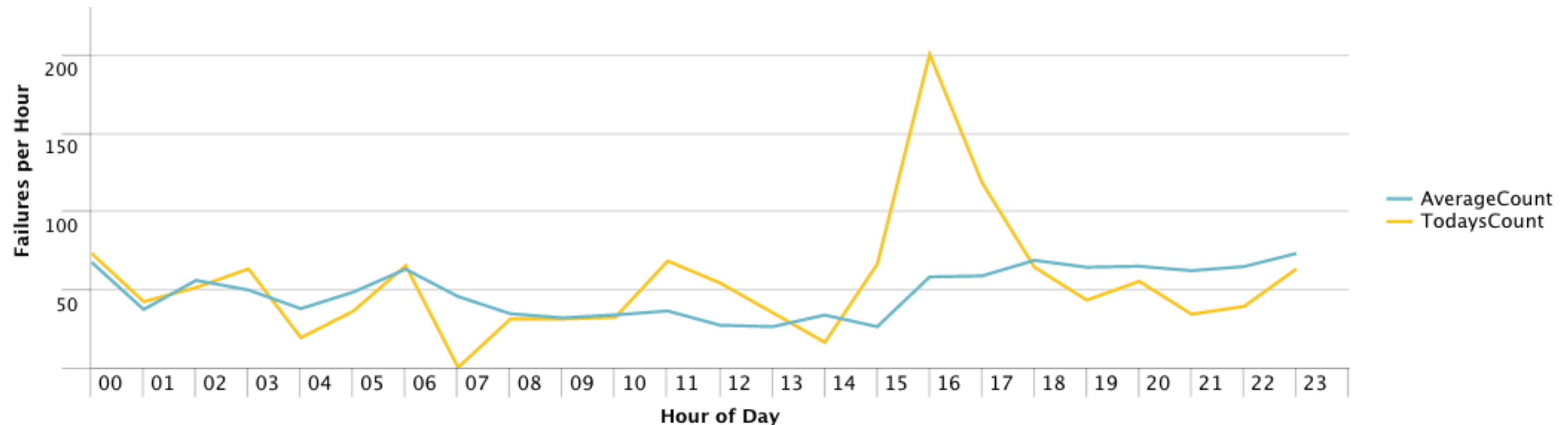
.conf2013

**YOUR DATA
NO LIMITS**

Computing "Across Time"

splunk>

Making Comparisons “Across Time”



AverageCount is the average number of failures *for that hour* across the last 30 days
TodaysCount is the number of failures for that hour in the last day

Average Events By Hour of Day Over Last 90 Days

? What is the average number of failures for each hour? We should end up with 24 averages, one for each hour of the day

1. Count the number of failures per hour of every day
2. Average the daily counts for the hour (average Monday at 1:00 with Tuesday at 1:00, etc.)


Average Events By Hour of Day Over Last 30 Days

? What is the average number of failures for each hour?

🔍 `tag=failure earliest=-30d latest=@d
| timechart span=1h count
| eval Hour = strftime(_time, "%H")
| stats avg(count) as AverageCount by Hour`

Compare the Last 24 Hours With the Average of the Last 30 Days

- Now that we have the average from the last search, how do we compare it with what is happening today?
- We will create a **subsearch** to search for today's data

 [**search** tag=failure earliest=-24h@h latest=@h
| timechart span=1h count as TodaysCount
| eval Hour = strftime(_time, "%H")]

Joining the Two Searches

🔍 tag=failure earliest=-30d latest=@d
| timechart span=1h count
| eval Hour = strftime(_time, "%H")
| stats avg(count) as AverageCount by Hour
| join Hour
 [search tag=failure earliest=-24h@h latest=@h
 | timechart span=1h count as TodaysCount
 | eval Hour = strftime(_time, "%H")]

Limitations of Subsearches

- Subsearches
 - Return only 100 results by default
 - Return 10,500 maximum results
 - Traverse the data independently of the main search
- Alternative:
 - Traverse the data only once for better performance
 - Avoid the result limits
 - Use eval command to categorize the event for the calculations

Eliminating the Join is Not Hard...

- Compare the last 24 hours with the average of the last 30 days

🔍 tag=failure earliest=-30d latest=@h
| timechart span=1h count
| eval StartTime=relative_time(now(),"-24h@h")
| eval Series=if(_time>=StartTime, "TodaysCount",
"AverageCount")
| eval Hour = strftime(_time, "%H")
| chart avg(count) by Hour Series

.conf2013

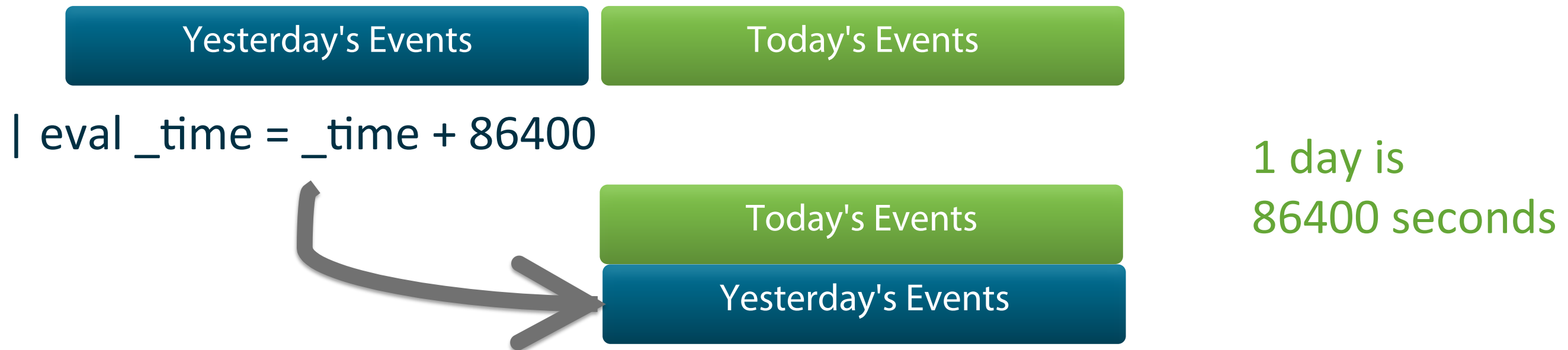
**YOUR DATA
NO LIMITS**

Comparing Equal Time Ranges

splunk>

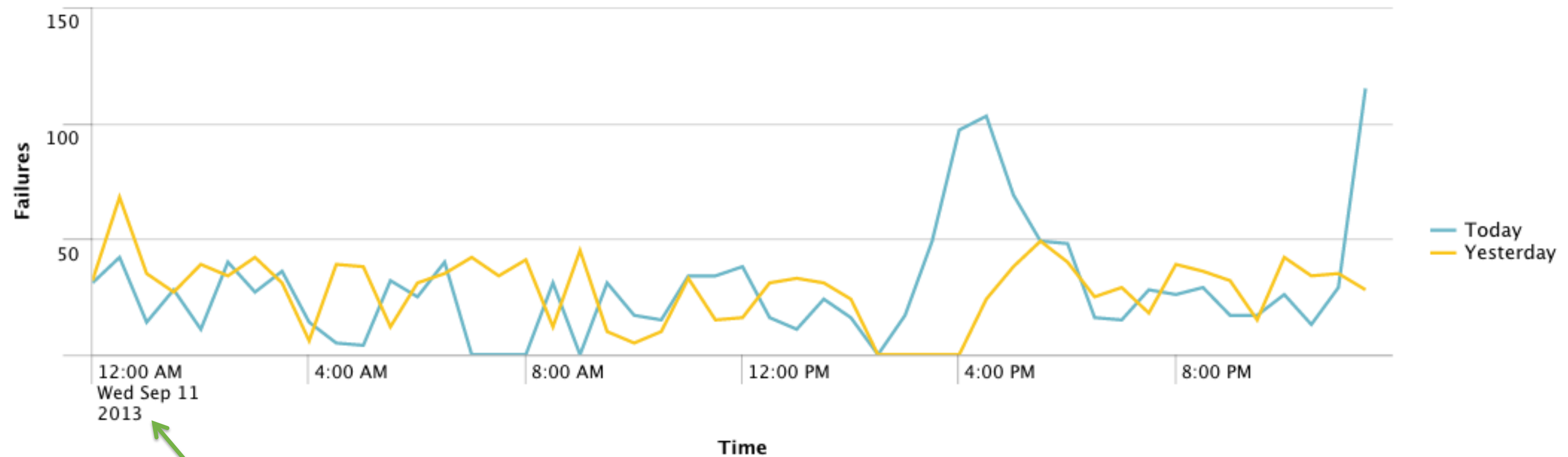
Comparing Equal Time Ranges

- When the time ranges are the same size
 - You can "slide" one time range over the other
 - Re-compute the time for each event




<http://blogs.splunk.com/2012/02/19/compare-two-time-ranges-in-one-report/>

Compare Yesterday and Today



all events appear as if they occurred today

How to Compare

 `tag=failure earliest=-1d@d latest=@d` outer search
| `eval Series="Yesterday"` retrieves yesterday's events,
| `eval _time = _time + 86400` labels them and
| `append [search tag=failure earliest=@d` recalculates the time
`latest=now`
| `eval Series = "Today"]`
| `timechart fixedrange=f span=30m count by Series`

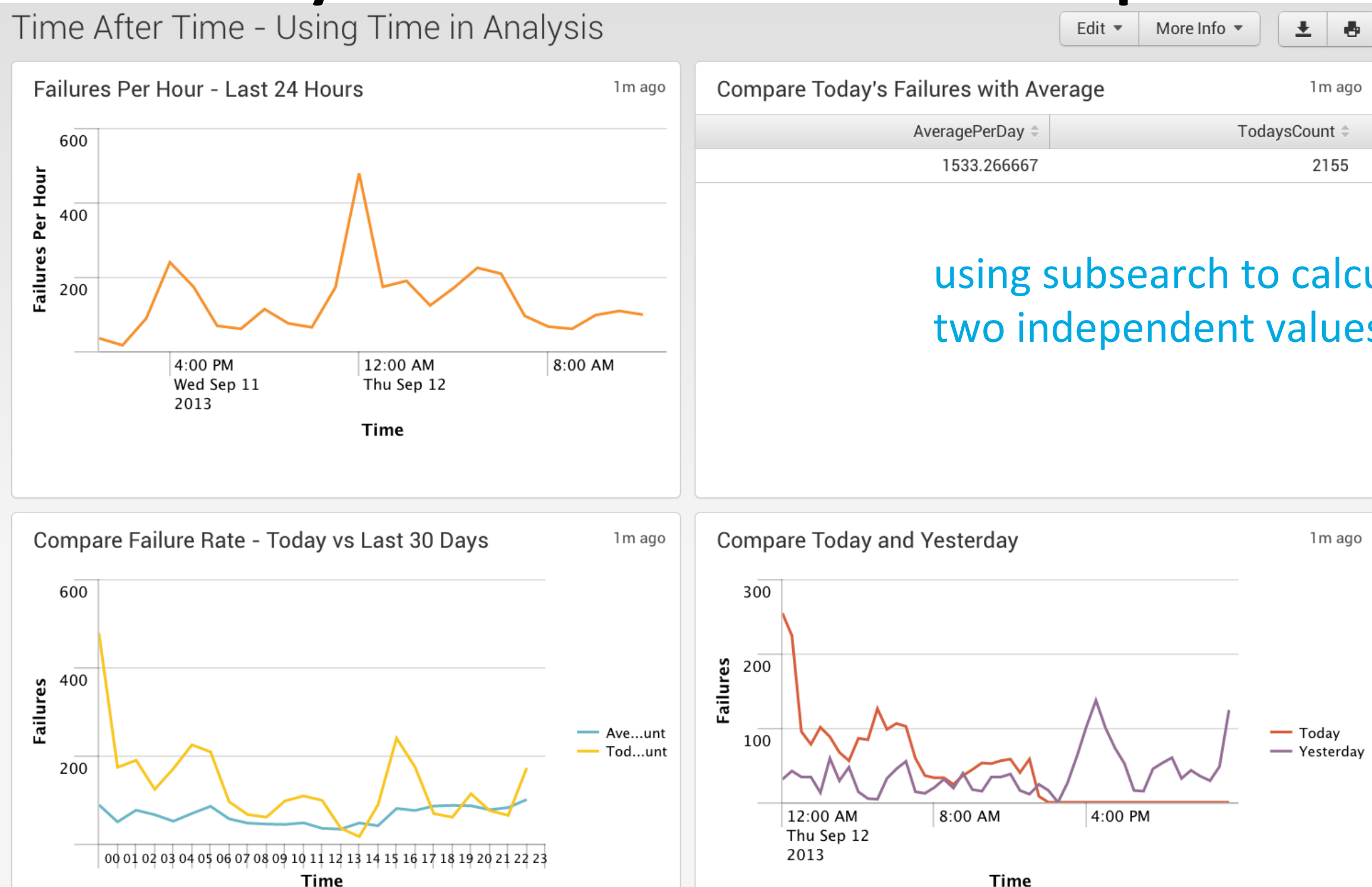
subsearch retrieves today's events
and labels them

When Time Ranges Are Equal

- There is no need to compute the statistics before combining the events
- The X-axis of the chart can easily be the timestamp instead of a calculated value

Summary: Time-based Comparisons

simple
timechart



comparing
different time
ranges

comparing
equal time
ranges

Summary

- Time can be manipulated like any other field
- Splunk provides time arithmetic and functions
- In addition to timechart, the chart and stats commands can be used for time-based data
- You can combine data and charts across time ranges, creating easy-to-understand visual comparisons

Next Steps

1

Download the .conf2013 Mobile App

If not iPhone, iPad or Android, use the Web App

2

Take the survey & **WIN A PASS FOR .CONF2014...** Or one of these bags!



.conf2013

**YOUR DATA
NO LIMITS**

THANK YOU

splunk>